INFORMATION CAPACITY OF THE STATIONARY GAUSSIAN CHANNEL

C.R. Baker* and S. Ihara**

LISS 39

July 1989

Abstract

The information capacity of the mismatched stationary continuous-time Gaussian channel is determined. The assumptions placed on the signal process are less restrictive than those given in previous treatments. Moreover, the assumptions used may have operational advantages over those used previously.

*Department of Statistics, University of North Carolina, Chapel Hill, N.C. 27599-3260

**Department of Mathematics, Faculty of General Education, Nagoya University, Chikusa-Ku, Nagoya 464, Japan

Research supported by ONR Grant N00014-89-J-1175 and NSF Grant NCR-8713727.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY 2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for Public Release: Distribution Unlimited
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Department of Statistics 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code) University of North Carolina Chapel Hill, North Carolina 27514	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-86-K-0039 and N00014-89-J-1175
8c. ADDRESS (City, State and ZIP Code) Statistics & Probability Program Arlington, VA 22217	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT NO.

11. TITLE (Include Security Classification)
Information Capacity of the Stationary Gaussian Channel

12. PERSONAL AUTHOR(S)
C.R. Baker and S. Ihara

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
TECHNICAL	FROM ______ TO ______	July 1989	33

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	Information capacity, Gaussian channel.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The information capacity of the mismatched stationary continuous-time Gaussian channel is determined. The assumptions placed on the signal process are less restrictive than those given in previous treatments. Moreover, the assumptions used may have operational advantages over those used previously.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED XX SAME AS RPT. XX DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL C.R. Baker	22b. TELEPHONE NUMBER (Include Area Code) (919) 962-2189 22c. OFFICE SYMBOL

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE. UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

Information Capacity of the Stationary Gaussian Channel

I. INTRODUCTION

Information capacity of the continuous-time stationary Gaussian channel (per unit time, as transmission time $T \to \infty$) has been a problem of much interest since the early days of information theory. Results date back to Shannon's 1949 paper [12], which contains an analysis of the band-limited channel. The treatment given in Gallager's 1968 book [6] is commonly referenced as "the" solution to the capacity problem. In fact, however, the class of channels that fit the model of [6] is quite limited, due to the nature of the constraints that are applied. These constraints eliminate from consideration a very large class of channels that can be important for applications.

The present paper contains an analysis of the capacity problem that applies to those channels not fitting the model used in Gallager's book [6]. The development is also quite different from that of [6]; it is based on the results of [3]. It will be seen that the treatment given here provides a desirable generality on the class of transmitted signals which is not present in previous treatments.

The model of the SGC (stationary Gaussian channel) is given by

$$Y(t) = X(t) + Z(t), \quad t \in \mathbb{R} = (-\infty,\infty), \tag{1}$$

where $X = \{X(t)\}$, $Y = \{Y(t)\}$ and $Z = \{Z(t)\}$ represent the channel input, output, and noise, respectively, Z is a stationary Gaussian process, and X is a stochastic process independent of Z. All stochastic processes mentioned in this paper will be assumed to be measurable and to have zero mean.

Shannon [12] (see also Fano [5, Chap. 5] and Pinsker [10]) studied the capacity per second of the SGC, defined by

$$\overline{C}(\mathcal{S}) = \sup_{X\in\mathcal{S}} \lim_{T\to\infty} \frac{1}{T} I_T(X,Y),$$

where $I_T(X,Y) = I(X_0^T, Y_0^T)$ denotes the mutual information between the input $X_0^T = \{X(t): 0 \le t < T\}$ and the corresponding output Y_0^T, and the supremum is taken for all admissible inputs $X = \{X(t)\} \in \mathcal{S}$. The class $\mathcal{S}$ of admissible inputs is given as follows. Given a stationary Gaussian process $W = \{W(t)\}$ with a spectral density function (SDF) f_W and a constant $P > 0$, an admissible input is a wide-sense stationary process X having a SDF f_X satisfying

$$\int_{-\infty}^{\infty} \frac{f_X(\lambda)}{f_W(\lambda)} d\lambda \le 2\pi P.$$

Formulas for $\overline{C}(\mathcal{S})$ have been given for special cases ([5],[10],[12]). However, it is known that there are some mathematical difficulties to derive the formulas. In this paper we shall derive the capacity (see (16) and (22)) in a rigorous way, assuming appropriate conditions on the processes W and Z. These assumptions imply, in particular, that the reproducing kernel Hilbert space (RKHS) of the process W is the same as the RKHS of Z.

In some previous works ([10],[12]) the stationarity of inputs is assumed. To remove the assumption of stationarity, we introduce a class $\mathfrak{X}$ of admissible inputs which we show to contain $\mathcal{S}$ under rather mild assumptions. Denote by $\|\cdot\|_{W,T}$ the norm of the RKHS $H_{W,T}$ corresponding to the process $W_0^T = \{W(t): 0 \le t < T\}$. The class $\mathfrak{X}$ is the set of all processes X satisfying

$$\overline{\lim_{T\to\infty}} \frac{1}{T} E[\|X_0^T\|_{W,T}^2] \le P. \tag{2}$$

Under modest assumptions, it will be shown that the capacity $\overline{C}(\mathfrak{X})$, defined in

the same manner as $\overline{C}(\mathcal{S})$, is equal to $\overline{C}(\mathcal{S})$. This means that, even if we do not require the stationarity of inputs, the capacity is attained in the class of stationary inputs.

A constraint similar to (2) has been considered by Baker [2],[3]. He determined the capacity

$$C^T = \sup_{X_0^T} \frac{1}{T} I_T(X,Y).$$

for a fixed time T, where the supremum is taken for all processes X_0^T satisfying

$$E[\|X_0^T\|^2_{W,T}] \leq P. \tag{3}$$

This result for C^T is essential for the derivation of our main results.

The average power constraints on the input studied in some previous works are given in terms of the noise process Z. The constraints (2) and (3) are stated in terms of a process W which is different from Z. The SGC subject to (3) is called a mismatched channel ([2],[3]). On the other hand, the SGC is called a matched channel if W = Z. The capacity of the matched Gaussian channel for a finite time interval is known to be equal to $\frac{1}{2}$ P ([1], [7]) and is not increased by feedback for a large class of channels [7].

In [6, Chap. 8], Gallager gives a comprehensive treatment of this problem, which he attributes to Holsinger. As mentioned above, his setup and assumptions are different. Relations between the Holsinger-Gallager result and those obtained here will be discussed following the statement of our main results. It will be shown that (when f_W and f_Z are rational) the assumptions of [6] are much more restrictive than those given here.

The precise definitions of the capacities are given in Section II. In Section III we give the statements of our main results, with some necessary lemmas given in Section V, and proofs in Section VI. Section IV contains a

discussion of these results, and comparisons to previous work.

When both f_W and f_Z are rational, the results given here and those of [6] provide a complete solution to the information capacity problem.

II. PRELIMINARIES

A basic assumption to be used throughout this paper is that the constraint is given in terms of a covariance function r_W defined by a spectral density function f_W. Moreover, the Gaussian noise process Z is assumed to be stationary with spectral density function f_Z and associated covariance function r_Z. The integral operator in $L_2[0,T]$ with kernel r_Z (resp., r_W) is denoted by $R_{Z,T}$ (resp., $R_{W,T}$). Since r_Z and r_W are both continuous, the reproducing kernel Hilbert space (RKHS) defined by each covariance function for the parameter set $[0,T]$ is isometric to the range of the (positive) square root of the associated integral operator in $L_2[0,T]$ (see, e.g., [4a]); we use the two spaces interchangeably, noting that the RKHS is a space of functions, while the range of the square root of the associated integral operator is a space of equivalence classes. It should be noted, however, that since we work with probabilities on the Borel sets of $L_2[0,T]$, letting $T \to \infty$, the range of the square root operator is the mathematically-correct definition. In order to have a finite value of $E\|X\|^2_{W,T}$, it is necessary [1] that $R_{X,T} = R^{\frac{1}{2}}_{W,T} L_T R^{\frac{1}{2}}_{W,T}$, where L_T is a trace-class operator with eigenvalues $\{\tau_n(T),\ n \geq 1\}$, and then

$$E\|X\|^2_{W,T} = \text{Trace } L_T = \sum_{n\geq 1} \tau_n(T). \tag{4}$$

We shall also assume that the covariance function r_Z of the stationary noise process Z is given by a spectral density function f_Z:

$$r_Z(t,s) \triangleq EZ(t)Z(s) = \int_{-\infty}^{\infty} e^{i\lambda(t-s)} f_Z(\lambda)\,d\lambda.$$

In this paper we consider the following classes of admissible input signals.

Definition 1 (Constraint on inputs):

(1) $\mathfrak{X}^T \equiv \mathfrak{X}_W^T(P)$ $(T > 0,\ P > 0)$ is the set of all $X_0^T = \{X(t): 0 \leq t < T\}$ such that the sample paths of X_0^T belong to $H_{W,T}$ with probability one and

$$E[\|X_0^T\|_{W,T}^2] \leq PT. \tag{5}$$

(2) $\mathfrak{X} \equiv \mathfrak{X}_W(P)$ is the set of all $X = \{X(t)\}$ such that $X_0^T \in H_{W,T}$ for all $T > 0$ with probability one and

$$\overline{\lim_{T\to\infty}} \frac{1}{T} E[\|X_0^T\|_{W,T}^2] \leq P \tag{6}$$

(3) $\mathcal{S} \equiv \mathcal{S}_W(P)$ is the set of all wide-sense-stationary processes $X = \{X(t)\}$ with SDF f_X such that f_X/f_W is bounded and

$$\int_{-\infty}^{\infty} \frac{f_X(\lambda)}{f_W(\lambda)}\, d\lambda \leq 2\pi P. \tag{7}$$

Let ξ and η be random variables or stochastic processes with paths in an appropriate space and denote by μ_ξ and $\mu_{\xi\eta}$ the induced probability distribution of ξ and the joint distribution of ξ and η, respectively. The mutual information $I(\xi,\eta)$ between ξ and η is defined by

$$I(\xi,\eta) \triangleq \int \log \frac{d\mu_{\xi\eta}}{d\mu_\xi \times \mu_\eta}\, d\mu_{\xi\eta},$$

if $\mu_{\xi\eta}$ is absolutely continuous with respect to the product measure $\mu_\xi \times \mu_\eta$, where $d\mu_{\xi\eta}/d\mu_\xi \times \mu_\eta$ is the Radon-Nikodym derivative; otherwise $I(\xi,\eta)$ is infinite. Denote by $I_T(X,Y) \triangleq I(X_0^T, Y_0^T)$ the mutual information between the input X_0^T and the corresponding output Y_0^T of the SGC, and define the mutual

information $\overline{I}(X,Y)$ per second by

$$\overline{I}(X,Y) \triangleq \lim_{T\to\infty} \frac{1}{T} I_T(X,Y),$$

if the limit exists. In order to be consistent with past work, particularly [1] - [3], the induced measures are defined on $L_2[0,T]$ for each fixed T.

The definition of capacity used in this paper is the mutual information version. Thus, the capacity of the SGC is the supremum of $I_T(X,Y)$ or $\overline{I}(X,Y)$ taken over all admissible inputs X. As is well-known, the information capacity is an upper bound on coding capacity.

Definition 2 (Capacity of SGC):

(1) $C^T \equiv C_W^T(P) \triangleq \sup \{\frac{1}{T} I_T(X,Y) : X_0^T \in \mathfrak{X}_W^T(P)\}$, $T > 0$.

(2) $\overline{C} \equiv \overline{C}_W(P) \triangleq \overline{\lim_{T\to\infty}} C_W^T(P)$.

(3) $\overline{C}(\mathfrak{X}) \equiv \overline{C}_W(P;\mathfrak{X}) \triangleq \sup \{\overline{\lim_{T\to\infty}} \frac{1}{T} I_T(X,Y) : X \in \mathfrak{X}_W(P)\}$.

(4) $\overline{C}(\mathcal{S}) \equiv \overline{C}_W(P;\mathcal{S}) \triangleq \sup \{\overline{\lim_{T\to\infty}} \frac{1}{T} I_T(X,Y) : X \in \mathcal{S}_W(P)\}$.

The capacity C^T has been determined by Baker ([2],[3]) for a general Gaussian channel, where W and Z are not necessarily stationary.

In order to state and prove our main results, we need some properties of mutual information.

Lemma 1 (see [1]): Suppose that the input X of SGC (1) is a Gaussian process. Then following conditions (a) - (c) are equivalent.

(a) $I_T(X,Y) < \infty$.

(b) The sample paths of X_0^T belong to $H_{Z,T}$ with probability one.

(c) The operator $R_{Z,T}^{-\frac{1}{2}} R_{X,T} R_{Z,T}^{-\frac{1}{2}}$ is of trace class.

If these conditions are satisfied, then $I_T(X,Y)$ is given by

$$I_T(X,Y) = \frac{1}{2} \sum_{n=1}^{\infty} \log(1 + \tau_n(T)),$$

where $\tau_n(T)$, $n = 1,2,\ldots$ are eigenvalues of $R_{Z,T}^{-\frac{1}{2}} R_{X,T} R_{Z,T}^{-\frac{1}{2}}$.

In the case where the input X is also a stationary Gaussian process, there have been various works on the mutual information (e.g., [5],[9],[11],[12]). A "well-known" formula of the mutual information per second is

$$\bar{I}(X,Y) = \frac{1}{4\pi} \int_{-\infty}^{\infty} \log\left[1 + \frac{f_X(\lambda)}{f_Z(\lambda)}\right] d\lambda. \tag{8}$$

It is known that there are some mathematical difficulties to derive (8) rigorously. Pinsker [11] proved (8) under some assumptions. Here, for later use, we restate some of Pinsker's results related to (8). We note by $\mathbb{F}$ the set of all SDF's $f(\lambda)$ which have the form $f = p(1 - \varphi)$ with a rational SDF p and a measurable function φ such that $0 \leq \varphi(\lambda) < 1$ for all $\lambda \in \mathbb{R}$ and $\int_{-\infty}^{\infty} |\log(1-\varphi(\lambda))| d\lambda < \infty$. A stationary Gaussian process is said to be information regular if $\lim_{T\to\infty} I(X_{-\infty}^{0}, X_{T}^{\infty}) = 0$. Necessary and sufficient conditions for information regularity are given in [8]. For example, if the SDF of X is rational, then X is information regular.

Lemma 2 [11 (Chap. 10), 9]: Let the input X be a stationary Gaussian process with SDF $f_X(\lambda)$. Then

(a) $\varliminf_{T\to\infty} \frac{1}{T} I_T(X,Y) \geq \frac{1}{4\pi} \int_{-\infty}^{\infty} \log\left[1 + \frac{f_X(\lambda)}{f_Z(\lambda)}\right] d\lambda$;

(b) If $f_X \in \mathbb{F}$ or $f_Y \in \mathbb{F}$, then (8) is true;

(c) If X or Y is information regular, then (8) is true.

For later use, we denote by $\mathbf{A}$ the class of all pairs (X,Z) of input X and noise Z for which (8) holds.

It is known (e.g., [1]) that if X and $\tilde{X} = \{\tilde{X}(t)\}$ have the same covariance function and X is Gaussian then $I_T(\tilde{X},\tilde{Y}) \leq I_T(X,Y)$, where $\tilde{Y}$ is the output corresponding to $\tilde{X}$. Noting (4) we know that the constraints (5) and (6) involve only covariance functions. Thus in order to find the capacity of the SGC we consider WLOG only Gaussian input signals.

We now consider the SGC (1) under the various constraints given in the preceding section.

Define a constant θ by

$$1 + \theta \triangleq \inf \{a: \int_{-\infty}^{\infty} \left[a - \min\left[a, \frac{f_Z(\lambda)}{f_W(\lambda)}\right]\right] d\lambda = \infty\}. \tag{9}$$

Roughly speaking, $1 + \theta = \lim_{\lambda\to\infty} f_Z(\lambda)/f_W(\lambda)$ if $f_W(\lambda)$ and $f_Z(\lambda)$ are smooth functions. The SDF f_Z of the noise process Z can be written in the form

$$f_Z = (1+\theta)f_W(1 + \sigma) = (1+\theta)f_W(1 + \sigma_+)(1 + \sigma_-), \tag{10}$$

where $\sigma_+(\lambda) \triangleq \max(\sigma(\lambda),0)$ and $\sigma_-(\lambda) \triangleq \min(\sigma(\lambda),0)$. We assume that $-1 < \theta < \infty$ and

$$1 + \sigma_-(\lambda) \geq a_0, \quad \lambda \in \mathbb{R}, \tag{11}$$

where $a_0 > 0$ is a constant.

For ease of notation, σ_+ and σ_- will be used in the proofs. However, in the statements of our main results σ_- will be replaced, using the following equality:

$$\begin{aligned} \sigma_-(\lambda) &= \frac{f_Z}{f_W}(\lambda)\,\frac{1}{1+\theta} - 1 & \qquad & \lambda \in \Lambda \\ &= 0 & & \text{otherwise} \end{aligned}$$

where $\Lambda = \{\lambda: f_Z(\lambda) < (1+\theta)f_W(\lambda)\}$. Note that (11) is equivalent to $(f_Z/f_W)(\lambda) \geq r_{\Lambda}(1+\theta)$.

III. MAIN RESULTS

We introduce auxiliary stationary Gaussian processes $V = \{V(t)\}$, $U = \{U(t)\}$ and $U^{(\epsilon)} = \{U^{(\epsilon)}(t)\}$ ($\epsilon > 0$) as follows. The process V is independent of U and $U^{(\epsilon)}$, and these processes have SDF's

$$f_V(\lambda) = f_W(\lambda)[1 + \sigma_+(\lambda)], \tag{12}$$

$$f_U(\lambda) = -f_W(\lambda)[1 + \sigma_+(\lambda)]\sigma_-(\lambda), \tag{13}$$

$$f_{U^{(\epsilon)}}(\lambda) = -f_W(\lambda)(1 + \sigma_+(\lambda))\sigma_-^{(\epsilon)}(\lambda),$$

respectively, where $\sigma_-^{(\epsilon)}(\lambda) \triangleq \sigma_-(\lambda)\cdot 1_{[-1,-\epsilon]}(\lambda)$ for $\lambda \in \mathbb{R}$ with the indicator function $1_{[-1,-\epsilon]}$ of a set $[-1,-\epsilon]$. Note that $f_Z = \theta(f_V - f_U)$ and consequently

$$R_{Z,T} = \theta(R_{V,T} - R_{U,T}). \tag{14}$$

If the formula (8) is valid for all Gaussian inputs $X \in \mathcal{S}$, then the capacity $\overline{C}(\mathcal{S})$ would be equal to

$$\Gamma(P) \triangleq \sup_f \left\{\frac{1}{4\pi}\int_{-\infty}^{\infty} \log\left[1 + \frac{f(\lambda)}{f_Z(\lambda)}\right]d\lambda\right\}, \tag{15}$$

where the supremum is taken for all SDF's f satisfying

$$\int_{-\infty}^{\infty} \frac{f(\lambda)}{f_W(\lambda)}\, d\lambda \leq 2\pi P.$$

Using the so-called water filling method (cf. [5],[6]) we can determine $\Gamma(P)$.

<u>Lemma 3</u>: Let $\Lambda = \{\lambda: f_Z(\lambda) < (1+\theta)f_W(\lambda)\}$.

I. If $P \geq \frac{1}{2\pi}\int_\Lambda\left[1 + \theta - \frac{f_Z}{f_W}(\lambda)\right]d\lambda$, then

$$\Gamma(P) = \frac{1}{4\pi}\int_{\Lambda}\log\left[(1+\theta)\frac{f_W}{f_Z}(\lambda)\right]d\lambda + \frac{P}{2(1+\theta)} + \frac{1}{4\pi(1+\theta)}\int_{\Lambda}\left[\frac{f_Z}{f_W}(\lambda) - (1+\theta)\right]d\lambda. \quad (16a)$$

II. If $P \leq \frac{1}{2\pi}\int_{\Lambda}\left[1 + \theta - \frac{f_Z}{f_W}(\lambda)\right]d\lambda$, then

$$\Gamma(P) = \frac{1}{4\pi}\int_{\Lambda_1}\log\left[(1+\theta)A(P)\frac{f_W}{f_Z}(\lambda)\right]d\lambda \quad (16b)$$

where $\Lambda_1 = \{\lambda\colon f_Z(\lambda) \leq (1+\theta)A(P)f_W(\lambda)\}$ and $A(P) \leq 1$ is uniquely defined by

$$P = \frac{1}{2\pi}\int_{\Lambda_1}\left[(1+\theta)A(P) - \frac{f_Z}{f_W}(\lambda)\right]d\lambda. \quad (17)$$

In Part I, $\Gamma(P) \geq P/[2(1+\theta)]$. In Part II, $\Gamma(P) \geq P/[2(1+\theta)A(P)] \geq P/2(1+\theta)$.

Theorem 1: (a) Assume that

$$\int_{\Lambda}\left[1 + \theta - \frac{f_Z}{f_W}(\lambda)\right]d\lambda < \infty. \quad (19)$$

If

$$\mathcal{S}_W(P) \subset \mathcal{R}_W(P) \quad (20)$$

and

$$(\alpha U, V) \in \quad \text{for sufficiently small } \alpha > 0 \quad (21)$$

are fulfilled, where $\alpha U = \{\alpha U(t)\}$, then all capacities per second coincide and are equal to $\Gamma(P)$:

$$\overline{C}_W(P;\mathcal{S}) = \overline{C}_W(P;\mathcal{R}) = \overline{C}_W(P) = \overline{C}_V(P) = \Gamma(P). \quad (22)$$

(b) Assume that

$$\int_{\Lambda}\left[1 + \theta - \frac{f_Z}{f_W}(\lambda)\right]d\lambda < \infty. \quad (23)$$

If (20) is satisfied and

$$(\alpha U^{(\epsilon)}, V) \in A \quad \text{for sufficiently small } \epsilon > 0 \text{ and } \alpha > 0, \quad (24)$$

then (22) holds.

Theorem 1 is derived from the following theorem.

<u>Theorem 2</u>: Following (C.1) - (C.5') holds.

(C.1) $\Gamma(P) \leq \overline{C}_W(P;\mathcal{S})$.

(C.2) If (20) is satisfied then

$$\overline{C}_W(P;\mathcal{S}) \leq \overline{C}_W(P;\mathfrak{A}).$$

(C.3) $\overline{C}_W(P;\mathfrak{A}) \leq \overline{C}_W(P+\epsilon), \quad \epsilon > 0.$

(C.3') If $\overline{C}_W(P)$ is continuous in P, then

$$\overline{C}_W(P;\mathfrak{A}) \leq \overline{C}_W(P).$$

(C.4) $\overline{C}_W(P) \leq \overline{C}_V(P)$.

(C.5) Under condition (19), if (21) is satisfied, then

$$\overline{C}_V(P) = \Gamma(P). \tag{25}$$

(C.5') Under condition (23), if (24) is satisfied, then (25) holds.

As one can see, (20) and (21) are key conditions for our assertion (22). Note that condition (20) is equivalent to

$$\lim_{T\to\infty} \frac{1}{2\pi T} E[\|X_0^T\|^2_{W,T}] \leq \int_{-\infty}^{\infty} \frac{f_X(\lambda)}{f_W(\lambda)}\, d\lambda, \quad X \in \mathcal{S}_W(P). \tag{26}$$

If (8) is true in general, then (21) holds. Although we do not have a proof, it is expected that (26) and (8) (or (20) and (21)) can be shown, so that (22) can be true under rather moderate conditions. We can give some sufficient conditions for (20) and (21).

<u>Theorem 3</u>: I. Suppose that (19) is satisfied.

(a) If $f_W \in \mathbb{F}$, then (20) is true.

(b) In addition to $f_W \in \mathbb{F}$, let $\frac{f_W}{f_Z}$ be continuous and $\frac{f_W}{f_Z}(\lambda) - (1+\theta) = O(\lambda^{-2})$ as $\lambda \to \infty$. Then (21) is true and (22) holds.

II. If f_W and f_Z are rational, then (19) is satisfied and (22) holds.

Moreover, related to (20), the following theorem may be of interest.

Theorem 4: Every stationary process in $\mathfrak{X}_W(P)$ belongs to $\mathcal{S}_W(P)$.

IV. DISCUSSION; RESULTS FOR RATIONAL SPECTRAL DENSITIES

In this section, we compare the results of Section III with those given in Section 8.5 of [6]. In order to clarify the comparison, we shall assume that f_W and f_Z are rational. Equations (19), (20), (21), and (22) then hold, according to Theorem 3, and all capacities are equal to $\Gamma(P)$ as given by (16).

The model used in [6] is as follows. The initial message process X_1 satisfies $E\int_0^T X_1^2(t)dt \leq PT$ for every $T > 0$. The transmitted signal, X, is obtained by passing X_1 through a linear filter with transfer function H. The assumption is that $|H(\lambda)|^2/f_Z < c_0 < \infty$, all $\lambda \in \mathbb{R}$, and $\int_{-\infty}^{\infty} \frac{|H(\lambda)|^2}{f_Z(\lambda)} d\lambda < \infty$. The capacity is defined as $\lim_{T\to\infty} C_T(P)/T$, where $C_T(P)$ is the capacity for the observation interval $[0,T]$.

To put this model into our context, we first note that $f_W(\lambda) = |H(\lambda)|^2$, $\lambda \in \mathbb{R}$. Since all capacities are equal when f_W and f_Z are rational, we can assume that X_1 is wide-sense stationary, with spectral density f_1, so that $f_X(\lambda) = f_1(\lambda)|H(\lambda)|^2$ for $\lambda \in \mathbb{R}$.

The assumption that $|H|^2/f_Z$ is bounded and integrable implies that the RKHS of f_W is a proper subset of the RKHS of f_Z. That is, y in $\mathcal{L}_2(0,\infty)$ belongs to the RKHS of f_W if and only if $\int_{-\infty}^{\infty} \frac{|\hat{y}(\lambda)|^2}{f_W(\lambda)} d\lambda < \infty$, where $\hat{y}$ is the L_2 Fourier

transform of y. Thus, $\int_{-\infty}^{\infty}\frac{|\hat{y}(\lambda)|^2}{f_Z(\lambda)}d\lambda = \int_{-\infty}^{\infty}\frac{|\hat{y}(\lambda)|^2}{f_W(\lambda)}\frac{f_W(\lambda)}{f_Z(\lambda)}d\lambda \le c_0\int_{-\infty}^{\infty}\frac{|\hat{y}(\lambda)|^2}{f_W(\lambda)}d\lambda$, so that y belongs to the RKHS of f_Z. The converse inclusion cannot hold; the function h with Fourier transform H belongs to the RKHS of Z, but not to that of W.

One of our assumptions is that $-1 < \theta < \infty$, where $1 + \theta$ is the infimum over all positive constants a satisfying

$$\int_{-\infty}^{\infty}\left[a - \min\left\{a, \frac{f_Z(\lambda)}{f_W(\lambda)}\right\}\right]d\lambda = \infty.$$

If f_W/f_Z is integrable and bounded, then $f_Z/f_W \to \infty$ as $|\lambda| \to \infty$, and so $\theta = \infty$. $\theta = \infty$ is equivalent to the RKHS of W being strictly contained in that of Z. If f_Z/f_W is integrable, then $\theta = -1$; this is equivalent to the RKHS of Z being strictly contained the RKHS of W. In fact, $-1 < \theta < \infty$ is equivalent to equality of the RKHS of W to that of Z; we omit the proof.

If the RKHS of W is not contained in that of Z, then $C_T(P) = \infty$ for every $T > 0$, by the results of [3]. Thus, the results of this paper and those of Section 8.5 of [6] together exhaust all cases for which the capacity is finite, and the problems treated do not overlap, when f_W and f_Z are rational.

From an operational viewpoint, the model used in this paper may be preferable to that used in [6]. That is, one might expect to attempt to limit the transmitted signal to some part of the noise band, defined by a linear filter with transfer function H_1. This would be done by filtering the received waveform. The effective noise then has spectral density $f_Z|H_1|^2$. The transmitted signal has effective spectral density $f_X|H_1|^2$. In order to provide the maximum amount of information, one would then wish to have the spectral density f_X be such that the signal sample paths, which are from the process with spectral density $f_X|H|^2$, use as much of the available bandwidth as possible. In essence, this means that the RKHS of $f_X|H|^2$ should equal that of $f_Z|H|^2$.

This would require that f_X and f_Z have the same RKHS. This cannot be achieved with the model of [6], whereas it is included in the model used here. It can be seen that the model of [6] limits the received signal to a relatively small part of the frequency region occupied by the effective noise process.

For example, if the effective noise at the receiver is described by a rational spectral density f_Z such that $f_Z(\lambda) \cong 1/\lambda^{2p}$ for $|\lambda| \to \infty$, then the model of [6] requires that $f_X(\lambda) \cong 1/\lambda^{4p}$ as $|\lambda| \to \infty$. The model used here permits $f_X(\lambda) \cong 1/\lambda^{2p}$ as $|\lambda| \to \infty$. In terms of signal sample paths, the sample paths of the signal process under the model of [6] would be required to be 4p-differentiable, while under the present assumptions, the signal sample paths need only be 2p-differentiable.

Thus, when f_W and f_Z are rational, we conclude the following:

(1) The problem treated in Section 8.5 of [6] and the problem treated here exhaust all situations where the capacity is finite.

(2) The problem treated here does not overlap with the problem treated in [6].

(3) The model treated here appears to have some advantages over the model of [6], from an operational viewpoint.

We now summarize our main results for the case where f_W and f_Z are rational, and include the main result given in [6] and the case $\theta = -1$ in order to cover all possible solutions.

<u>Theorem 5</u>. Suppose that f_W and f_Z are rational. Then the capacity with the input process required to be stationary is the same as the capacity without this restriction: $\overline{C}_W(P;\mathcal{S}) = \overline{C}_W(P;\mathfrak{A})$. For $-1 \leq \theta \leq \infty$, the value of the capacity is as follows.

1. $\theta = -1$. In this case $\bar{C}_W(P;\mathfrak{R}) = \infty$, $C_W^T(P) = \infty$ for every $T > \infty$, and this holds with the additional assumption that X_0^T is a one-dimensional process.

2. $-1 < \theta < \infty$. Let $\Lambda = \{\lambda: f_Z(\lambda) < (1+\theta)f_W(\lambda)\}$.

   I. If $P \geq \frac{1}{2\pi}\int_\Lambda\left[1 + \theta - \frac{f_Z}{f_W}(\lambda)\right]d\lambda$, then

$$\bar{C}_W(P;\mathfrak{R}) = \frac{1}{4\pi}\int_\Lambda \log\left[(1+\theta)\frac{f_W}{f_Z}(\lambda)\right]d\lambda + \frac{P}{2(1+\theta)} + \frac{1}{4\pi(1+\theta)}\int_\Lambda\left[\frac{f_Z}{f_W}(\lambda) - (1+\theta)\right]d\lambda.$$

   II. If $P \leq \frac{1}{2\pi}\int_\Lambda\left[1 + \theta - \frac{f_Z}{f_W}(\lambda)\right]d\lambda$, then

$$\bar{C}_W(P;\mathfrak{R}) = \frac{1}{4\pi}\int_{\Lambda_1}\log\left[(1+\theta)A(P)\frac{f_W}{f_Z}\right]d\lambda \qquad \text{where}$$

   $\Lambda_1 = \{\lambda: f_Z(\lambda) \leq (1+\theta)A(P)f_W(\lambda)\}$ and $A(P) \leq 1$ is uniquely defined by

$$P = \frac{1}{2\pi}\int_{\Lambda_1}\log\left[(1+\theta)A(P) - \frac{f_Z}{f_W}(\lambda)\right]d\lambda.$$

   In Part I, $\bar{C}_W(P;\mathfrak{R}) \geq P/[2(1+\theta)]$. In Part II, $\bar{C}_W(P;\mathfrak{R}) \geq P/[2(1+\theta)A(P)] \geq P/2(1+\theta)$.

3. $\theta = \infty$. Then [6]

$$\bar{C}_W(P;\mathfrak{R}) = \frac{1}{4\pi}\int_{\Lambda_2}\log\left[B(P)\frac{f_W}{f_Z}(\lambda)\right]d\lambda,$$

   where $\Lambda_2 = \{\lambda: f_Z(\lambda) \leq B(P)f_W(\lambda)\}$ and $B(P)$ is uniquely determined by

$$P = \frac{1}{2\pi}\int_{\Lambda_2}\left[B(P) - \frac{f_Z}{f_W}(\lambda)\right]d\lambda.$$

Moreover, in part II of 2) and in 3), capacity can be attained with a stationary Gaussian signal process X with spectral density f_X. For Part II of 2), f_X is defined by

$$f_X(\lambda) = (1+\theta)A(P) - [f_Z/f_W](\lambda) \qquad \lambda \text{ in } \Lambda_1$$
$$= 0 \qquad \text{otherwise.}$$

For 3), f_X is defined by

$$f_X(\lambda) = B(P) - [f_Z/f_W](\lambda) \qquad \lambda \text{ in } \Lambda_2$$
$$= 0 \qquad \text{otherwise}.$$

In considering Theorem 5, it can be seen that II of Part 2) is quite similar to the Holsinger-Gallager result given in Part 3); in fact, Part 3) can be formally obtained by substituting $B(P) \equiv (1+\theta)A(P)$ in II. To gain insight into this, one can compare these results with those contained in Theorem 3 and Corollary 4 of [3]. Specifically, Part I above should be compared with Part (a) of Theorem 3 in [3]; II of Part 2) above should be compared with Part (b) of Theorem 3 in [3]; and Part 3) should be compared with Corollary 4 of [3].

The question of attaining capacity (in particular, by a stationary process) in I of Part 2) is still open. For any finite T, $C_W^T(PT)$ cannot be attained when the conditions of Part I are satisfied (see Theorem 3(e) of [3]).

V. LEMMAS

In this section we shall give several lemmas needed to prove our main results. Proofs of the lemmas will be given in the Appendix.

Hereafter, for brevity, Σ_n signifies $\Sigma_{n=1}^{\infty}$ and $\int$ signifies $\int_{-\infty}^{\infty}$ unless noted otherwise.

We define an operator S_T on $L_2[0,T]$ by

$$S_T \triangleq \frac{1}{\theta} R_{V,T}^{-\frac{1}{2}} R_{Z,T} R_{V,T}^{-\frac{1}{2}} - I, \tag{27}$$

where I is the identity operator.

Lemma 4: Suppose that the condition (21) is satisfied. Then S_T is a negative definite trace class operator satisfying

$$S_T = -R_{V,T}^{-\frac{1}{2}} R_{U,T} R_{V,T}^{-\frac{1}{2}}. \tag{28}$$

The eigenvalues $\lambda_n(T)$, $n = 1,2,\ldots$, of S_T are bounded by

$$-1 + a_0 \leq \lambda_n(T) \leq 0, \quad n = 1,2,\ldots, \tag{29}$$

where a_0 is the constant of (11).

The capacity of a mismatched Gaussian channel is given in [3]. That result can be directly applied for our SGC. We arrange the eigenvalues of S_T in increasing order: $\lambda_1(T) \leq \lambda_2(T) \leq \ldots$.

<u>Lemma 5</u> ([2],[3]): Suppose that (19) and (21) are satisifed. Then $C_V^T(P)$ are given by

$$C_V^T(P) = \frac{1}{2T} \Sigma_n \log \left[(1+\lambda_n(T))^{-1}\right] + \frac{P}{2\theta} + \frac{1}{2T} \Sigma_n \lambda_n(T),$$
$$\text{if } PT \geq -\theta\, \Sigma_n \lambda_n(T), \tag{30a}$$

$$C_V^T(P) = \frac{1}{2T} \sum_{n=1}^{K} \log \left[\frac{\Sigma_{i=1}^{K} \lambda_i(T) + \theta^{-1} PT + K}{K(1+\lambda_n(T))}\right],$$
$$\text{if } PT < -\theta\, \Sigma_n \lambda_n(T), \tag{30b}$$

where $K \equiv K(T) \equiv K(T,P)$ is the largest integer such that

$$\sum_{i=1}^{K} \lambda_i(T) + \theta^{-1} PT \geq K\lambda_K(T). \tag{31}$$

<u>Remark</u>: We note that if the SGC is a matched one then $S_T = 0$ and (30) is reduced to

$$C_V^T(P) = P/(2\theta), \quad P > 0. \tag{32}$$

To examine the asymptotic behavior of the eigenvalues $(\lambda_n(T))$ as $T \to \infty$, we prepare a lemma. Let $(\tau_n(T))$ $(T > 0)$ be a summable sequence and $\tau(\lambda)$,

$\lambda \in \mathbb{R}$, be an integrable function such that

$$0 \le \tau_n(T) \le M, \quad 0 \le \tau(\lambda) \le M,$$

for all $n = 1,2,\ldots$, $T > 0$ and $\lambda \in \mathbb{R}$, where M is a constant. Let F be a bounded piecewise continuous function defined on $[0,M]$ such that F is continuous at $x = 0$ and $F(0) = 0$.

Lemma 6: (a) Suppose that $\{\tau_n(T)\}$ and τ satisfy

$$\lim_{T\to\infty} \frac{1}{T} \Sigma_n \log\,(1+\alpha\tau_n(T)) = \frac{1}{2\pi}\int \log\,(1+\alpha\tau(\lambda))d\lambda \tag{33}$$

for every $0 < \alpha \le \alpha_0$, where α_0 is a constant. Then it holds that

$$\lim_{T\to\infty} \frac{1}{T} \Sigma_n\, F(\tau_n(T)) = \frac{1}{2\pi}\int F(\tau(\lambda))d\lambda. \tag{34}$$

In particular,

$$\lim_{T\to\infty} \frac{1}{T} \Sigma_n\, \tau_n(T)^k = \frac{1}{2\pi}\int \tau(\lambda)^k d\lambda, \quad k = 1,2,\ldots. \tag{35}$$

(b) Suppose that

$$\underline{\lim}_{T\to\infty} \frac{1}{T} \Sigma_n \log(1 + \alpha\tau_n(T)) \ge \frac{1}{2\pi}\int \log(1 + \alpha\tau(\lambda))d\lambda$$

for every $0 < \alpha \le \alpha_0$. Then

$$\underline{\lim}_{T\to\infty} \frac{1}{T} \Sigma_n\, \tau_n(T)^k \ge \frac{1}{2\pi}\int \tau(\lambda)^k d\lambda, \quad k = 1,2,\ldots.$$

Lemma 7: For any $X \in \mathfrak{X}_W(P)$, it holds that

$$\text{Trace } R_{V,T}^{-\frac{1}{2}} R_{X,T} R_{V,T}^{-\frac{1}{2}} \le \text{Trace } R_{W,T}^{-\frac{1}{2}} R_{X,T} R_{W,T}^{-\frac{1}{2}}, \tag{36}$$

or equivalently

$$E[\|X_0^T\|^2_{V,T}] \le E[\|X_0^T\|^2_{W,T}]. \tag{37}$$

Lemma 8: Let f be a continuous SDF such that $f(\lambda) = O(\lambda^{-2})$ as $\lambda \to \infty$. Then for any $\epsilon > 0$ there exists a rational SDF $g(\lambda)$ such that $f(\lambda) \leq g(\lambda)$, $\lambda \in \mathbb{R}$, and

$$\int_{-\infty}^{\infty} [g(\lambda) - f(\lambda)]d\lambda < \epsilon.$$

VI. PROOF OF THEOREMS

At first we prove Theorem 2.

Proof of (C.1) - (C.4): (C.1) follows from (a) of Lemma 2. (C.2) is clear from the definition of the capacities. Let $X \in \mathfrak{X}_W(P)$. Then

$$\frac{1}{T} E[\|X_0^T\|_{W,T}^2] \leq P + \epsilon$$

for sufficiently large T. Therefore

$$\overline{\lim_{T\to\infty}} \frac{1}{T} I_T(X,Y) \leq C_W^T(P+\epsilon)$$

and we obtain (C.3). It is clear from Lemma 7 that $\mathfrak{X}_W^T(P) \subset \mathfrak{X}_V^T(P)$, yielding (C.4). (C.3') is clear from (C.3).

Proof of (C.5): For each $\alpha > 0$ we define a process $Q^{(\alpha)} = \{Q^{(\alpha)}(t)\}$ by $Q^{(\alpha)}(t) = \sqrt{\alpha}\, U(t) + V(t)$. Then by Lemma 1 and (28), we have

$$I_T(U,Q^{(\alpha)}) = \tfrac{1}{2} \Sigma_n \log(1 - \alpha\lambda_n(T)), \tag{38}$$

where $\lambda_n(T)$, $n = 1,2,\ldots$, are the eigenvalues of the operator S_T defined by (27). Hence, noting (12) and (13), the assumption (21) implies that

$$\lim_{T\to\infty} \frac{1}{T} \Sigma_n \log(1 - \alpha\lambda_n(T)) = \frac{1}{4\pi} \int \log(1 - \alpha\sigma_-(\lambda))d\lambda \tag{39}$$

for sufficiently small α. We can apply Lemma 6, replacing $\tau_n(T)$ and $\tau(\lambda)$ by $-\lambda_n(T)$ and $-\sigma_-(\lambda)$, respectively. At first we shall show (25) for

$P \geq \frac{-\theta}{2\pi} \int \sigma_-(\lambda)d\lambda$. It follows from (35) that

$$\lim_{T\to\infty} \frac{-\theta}{T} \Sigma_n \lambda_n(T) = \frac{-\theta}{2\pi} \int \sigma_-(\lambda)d\lambda \leq P. \tag{40}$$

Noting (11) and (29) we can apply (34) for $F(x) = \log(1-x)$ to obtain

$$\lim_{T\to\infty} \frac{1}{T} \Sigma_n \log(1 + \lambda_n(T)) = \frac{1}{2\pi} \int \log(1 + \sigma_-(\lambda))d\lambda. \tag{41}$$

Since $C_V^T(P)$ is given by (30), using (40) and (41), we obtain

$$\bar{C}_V(P) = \lim_{T\to\infty} C_V^T(P) = \lim_{T\to\infty} \frac{1}{T} \{\tfrac{1}{2} \Sigma_n \log[(1+\lambda_n(T))^{-1}] + \frac{PT}{2\theta} + \tfrac{1}{2} \Sigma_n \lambda_n(T)\}$$

$$= \frac{1}{4\pi} \int \log[(1+\sigma_-(\lambda))^{-1}]d\lambda + \frac{P}{2\theta} + \frac{1}{4\pi} \int \sigma_-(\lambda)d\lambda = \Gamma(P).$$

Secondarily we shall show (25) *for* $P < \frac{-\theta}{2\pi} \int \sigma_-(\lambda)d\lambda$. It will be shown that

$$\lim_{T\to\infty} \lambda_{K(T)}(T) = A(P), \tag{42}$$

where $K = K(T)$ is the constant in (30b) and $A(P)$ is the constant given by (17). To prove (42) we put

$$\bar{A} = \overline{\lim_{T\to\infty}} \lambda_{K(T)}(T), \quad \underline{A} = \underline{\lim}_{T\to\infty} \lambda_{K(T)}(T).$$

There exists $\{T_n\}$ such that $T_n \to \infty$ as $n \to \infty$ and

$$\lim_{n\to\infty} \lambda_{K(T_n)}(T_n) = \bar{A}.$$

Applying (34) for $F(x) = 1_{[-\bar{A},0]}(x)\cdot x$ ($1_B(x)$ is the indicator function of a set B) we obtain

$$\lim_{n\to\infty} \frac{1}{T_n} \sum_{k=1}^{K(T_n)} \lambda_k(T_n) = \lim_{n\to\infty} \frac{1}{T_n} \sum_{k;\lambda_k(T_n)\leq\bar{A}} \lambda_k(T_n) \tag{43}$$

$$= \frac{1}{2\pi} \int_{D(\bar{A})} \sigma_-(\lambda)d\lambda.$$

where $D(\bar{A}) = \{\lambda;\ \sigma_-(\lambda) \leq \bar{A}\}$. By definition we have

$$\sum_{i=1}^{K(T)+1} \lambda_i(T) + \frac{PT}{\theta} < (K(T)+1)\lambda_{K(T)+1}(T). \tag{44}$$

From (31), (43) and (44), we obtain

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{T_n} K(T_n)\lambda_{K(T_n)}(T_n) &= \lim_{n\to\infty} \frac{1}{T_n} \sum_{k=1}^{K(T_n)} \lambda_k(T_n) + \frac{P}{\theta} \\ &= \frac{1}{2\pi} \int_{D(\bar{A})} \sigma_-(\lambda)d\lambda + \frac{P}{\theta}.\end{aligned} \tag{45}$$

On the other hand, applying (34) for $F(x) = 1_{[-\bar{A},o]}(x)$ we get

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{T_n} K(T_n)\lambda_{K(T_n)}(T_n) &= \bar{A} \lim_{n\to\infty} \frac{1}{T_n} \#\{k;\ \lambda_k(T_n) \leq \bar{A}\} \\ &= \frac{1}{2\pi} \bar{A}|D(\bar{A})|,\end{aligned} \tag{46}$$

where #B denotes the cardinal number of a set B and $|D|$ denotes the Lebesgue measure of a set D. By (45) and (46) we know that $\bar{A}$ is a solution of (17). Since (17) has the unique solution A(P), $\bar{A}$ is equal to A(P). In the same way it is shown that $\underline{A}$ is also equal to A(P). Thus we get (42). In the same manner as above, applying (34), we obtain the following equations.

$$\lim_{T\to\infty} \frac{K(T)}{T} = \lim_{T\to\infty} \frac{1}{T} \#\{k;\ \lambda_k(T) \leq A(P)\} = \frac{1}{2\pi} |D(A(P))|, \tag{47}$$

$$\begin{aligned}\lim_{T\to\infty} \frac{1}{T} \sum_{n=1}^{K(T)} \log(1+\lambda_n(T)) &= \lim_{T\to\infty} \sum_{\lambda_n(T)\leq A(P)} \frac{1}{T} \log(1+\lambda_n(T)) \\ &= \frac{1}{2\pi} \int_{D(A(P))} \log(1+\sigma_-(\lambda))d\lambda,\end{aligned} \tag{48}$$

$$\lim_{T\to\infty} \frac{1}{T} \sum_{k=1}^{K(T)} \lambda_k(T) = \lim_{T\to\infty} \frac{1}{T} \sum_{\lambda_k(T)\leq A(P)} \lambda_k(T) = \frac{1}{2\pi} \int_{D(A(P))} \sigma_-(\lambda)d\lambda. \tag{49}$$

It follows from (30), (47) - (49) and (17) that

$$\bar{C}_V(P) = \lim_{T\to\infty} C_V^T(P)$$

$$= \lim_{T\to\infty} \frac{-1}{2T} \sum_{n=1}^{K(T)} \log(1+\lambda_n(T)) + \frac{|D(A(P))|}{4\pi} \lim_{T\to\infty} \log\left[1 + \frac{T}{K(T)}\left\{\frac{1}{T} \sum_{k=1}^{K(T)} \lambda_k(T) + \frac{P}{\theta}\right\}\right]$$

$$= \frac{-1}{4\pi} \int_{D(A(P))} \log(1+\sigma_-(\lambda))d\lambda$$

$$+ \frac{|D(A(P))|}{4\pi} \log\left[1 + \frac{1}{|D(A(P))|}\left\{\int_{D(A(P))} \sigma_-(\lambda)d\lambda + \frac{2\pi P}{\theta}\right\}\right]$$

$$= \frac{1}{4\pi} \int_{D(A(P))} \log\left[\frac{1 + A(P)}{1 + \sigma_-(\lambda)}\right]d\lambda = \Gamma(P).$$

Proof of (C.5'): Let $Z^{(\epsilon)} = \{Z^{(\epsilon)}(t)\}$ $(\epsilon > 0)$ be a stationary Gaussian process with $f^{(\epsilon)}(\lambda) = \theta f_V(\lambda)(1+\sigma_-^{(\epsilon)}(\lambda))$ as the SDF and consider an SGC

$$Y^{(\epsilon)}(t) = X(t) + Z^{(\epsilon)}(t), \quad t \in \mathbb{R}. \tag{50}$$

From (9) it is clear that $-\int\sigma_-^{(\epsilon)}(\lambda)d\lambda < \infty$. Therefore, we can apply (C.5) to get the capacity

$$\bar{C}_V^{(\epsilon)}(P) \triangleq \overline{\lim_{T\to\infty}} \frac{1}{T} \sup \{I_T(X,Y^{(\epsilon)});\ X_0^T \in \mathcal{X}_V^T(P)\} \tag{51}$$

of SGC (50). Noting (23) it can be shown for every fixed P that

$$\bar{C}_V^{(\epsilon)}(P) = \Gamma(P) \quad \text{for sufficiently small } \epsilon > 0. \tag{52}$$

Since

$$f^{(\epsilon)}(\lambda) \geq f_Z(\lambda),$$

using Lemma 1 we can show that

$$I_T(X,Y^{(\epsilon)}) \leq I_T(X,Y)$$

so that

$$\overline{C}_V^{(\epsilon)}(P) \le \overline{C}_V(P). \tag{53}$$

On the other hand it can be shown that

$$\lim_{\epsilon\to 0} I_T(X,Y^{(\epsilon)}) = I_T(X,Y).$$

And consequently, for every $\delta > 0$, we can show that there exists $\epsilon_0 > 0$ such that

$$\overline{C}_V(P) \le \overline{C}_V^{(\epsilon)}(P) + \delta, \quad 0 < \epsilon \le \epsilon_0. \tag{54}$$

The desired equation (25) follows from (52) - (54).

Proof of Theorem 1: (a) It follows from (C.1) - (C.5) that

$$\begin{aligned}\Gamma(P) &\le \overline{C}_W(P;\mathcal{G}) \le \overline{C}_W(P;\mathfrak{R}) \le \overline{C}_W(P+\epsilon) \\ &\le \overline{C}_V(P+\epsilon) = \Gamma(P+\epsilon), \quad \epsilon > 0.\end{aligned} \tag{55}$$

Since $\Gamma(P)$ is continuous in P, we obtain (22) from (55).

(b) can be shown in the same way.

Proof of Theorem 3: I. (a) Let $X \in \mathcal{G}_W(P)$. By definition $\tau(\lambda) \triangleq f_X(\lambda)/f_W(\lambda)$ is bounded ($0 \le \tau(\lambda) \le M < \infty$) and $\int\tau(\lambda)d\lambda \le 2\pi P$. Denote by $\tau_n(T)$, $n = 1,2,\ldots,$ the eigenvalues of $R_{W,T}^{-\frac{1}{2}}R_{X,T}R_{W,T}^{-\frac{1}{2}}$. Let $\xi = \{\xi(t)\}$ and $\zeta = \{\zeta(t)\}$ be mutually independent Gaussian stationary processes, which are independent of X and W, with SDF's $f_\xi = \alpha f_W\tau$ and $f_\zeta = f_W(1-\alpha\tau)$, where $0 < \alpha < 1/M$. And let f_η be the SDF of the output process $\eta = \{\eta(t)\}$ of an SGC

$$\eta(t) = \xi(t) + \zeta(t), \quad t \in \mathbb{R}.$$

Then it is clear that

$$f_\eta = f_\xi + f_\zeta = f_W \in \mathbb{F}.$$

Therefore, by (b) of Lemma 2, we have

$$\overline{I}(\xi,\eta) = \frac{1}{4\pi}\int\left[1 + \frac{\alpha\tau(\lambda)}{1-\alpha\tau(\lambda)}\right]d\lambda. \tag{56}$$

We consider another SGC

$$\theta(t) = \xi(t) + W(t).$$

Since $f_\zeta(\lambda) \le f_W(\lambda)$ for all $\lambda \in \mathbb{R}$, using Lemma 1 we can show that

$$I_T(\xi,\eta) \ge I_T(\xi,\theta) = \tfrac{1}{2}\,\Sigma_n \log(1 + \alpha\tau_n(T)). \tag{57}$$

Using (a) of Lemma 2, (56) and (57) we obtain

$$\begin{aligned}\frac{1}{4\pi}\int \log(1+\alpha\tau(\lambda))d\lambda &\le \varliminf_{T\to\infty} \frac{1}{2T}\,\Sigma_n \log(1+\alpha\tau_n(T)) \\ &\le \varlimsup_{T\to\infty} \frac{1}{2T}\,\Sigma_n \log(1+\alpha\tau_n(T)) \le \bar{I}(\xi,\eta) \\ &= \frac{1}{4\pi}\int \log\left[1 + \frac{\alpha\tau(\lambda)}{1-\alpha\tau(\lambda)}\right]d\lambda. \end{aligned} \tag{58}$$

Since $\alpha > 0$ is arbitrary, in the same way as Lemma 6, we can derive

$$\lim_{T\to\infty} \frac{1}{T}\,\Sigma_n \tau_n(T) \le \frac{1}{2\pi}\int \tau(\lambda)d\lambda \tag{59}$$

from (58). Noting (4) we know that (59) is equivalent to (26), and to (20).

(b) By Lemma 8, for any $\epsilon > 0$, there exists a rational SDF $g(\lambda)$ such that $0 \le -\sigma_- \le g(\lambda)$, $\lambda \in \mathbb{R}$, and

$$\int_{-\infty}^{\infty} [g(\lambda) + \sigma_-(\lambda)]d\lambda < \epsilon. \tag{1}$$

Let $\tilde{U} = \{\tilde{U}(t)\}$ be a Gaussian stationary process with SDF $f_W(\lambda)g(\lambda)$ and independent of the process V. Let $\alpha > 0$ be fixed and define processes $W_\alpha = \{W_\alpha(t)\}$ and $\tilde{W} = \{\tilde{W}_\alpha(t)\}$ by

$$W_\alpha(t) = \alpha U(t) + V(t) \quad \text{and} \quad \tilde{W}_\alpha(t) = \alpha\tilde{U}(t) + V(t).$$

Since g is rational, the SDF $f_W \cdot g$ of $\tilde{U}$ belongs to $\mathbb{F}$. Hence $(\alpha\tilde{U},V) \in A$ and

$$\bar{I}(\tilde{U},\tilde{W}_\alpha) = \frac{1}{4\pi}\int_{-\infty}^{\infty} \log\left[1 + \frac{\alpha^2 f_W g}{f_V}\right]d\lambda = \frac{1}{4\pi}\int_{-\infty}^{\infty} \log\left[1 + \frac{\alpha^2 g}{1+\sigma_+}\right]d\lambda. \tag{2}$$

Noting

$$f_U(\lambda) = -f_W(\lambda)(1+\sigma_+(\lambda))\sigma_-(\lambda) = -f_W(\lambda)\sigma_-(\lambda) \le f_W(\lambda)g(\lambda) = f_{\tilde{U}}(\lambda),$$

we can show that

$$I_T(U,W_\alpha) \le I_T(\tilde{U},\tilde{W}_\alpha), \quad T > 0. \tag{3}$$

Using Lemma 2 we have

$$\underline{\lim}_{T\to\infty} \frac{1}{T} I_T(U,W_\alpha) \ge \frac{1}{4\pi}\int_{-\infty}^{\infty} \log\left[1 + \frac{\alpha^2 f_U}{f_V}\right]d\lambda = \frac{1}{4\pi}\int_{-\infty}^{\infty} \log\left[1 - \frac{\alpha^2\sigma_-}{1+\sigma_+}\right]d\lambda. \tag{4}$$

It follows from (1) that

$$\frac{1}{4\pi}\int_{-\infty}^{\infty} \log\left[1 + \frac{\alpha^2 g}{1+\sigma_+}\right]d\lambda - \frac{1}{4\pi}\int_{-\infty}^{\infty} \log\left[1 - \frac{\alpha^2\sigma_-}{1+\sigma_+}\right]d\lambda$$

$$= \frac{1}{4\pi}\int_{-\infty}^{\infty} \log\left[1 + \frac{\alpha^2(g+\sigma_-)}{1+\sigma_+-\alpha^2\sigma_-}\right]d\lambda$$

$$\le \frac{1}{4\pi}\int_{-\infty}^{\infty} \log\left[1 + \alpha^2(g+\sigma_-)\right]d\lambda \le \frac{\alpha^2}{4\pi}\int_{-\infty}^{\infty} (g+\sigma_-)d\lambda < \frac{\sigma^2\epsilon}{4\pi}. \tag{5}$$

Using (2) - (5) we have

$$\overline{I}(\tilde{U},\tilde{W}_\alpha) \ge \overline{\lim}_{T\to\infty} \frac{1}{T} I_T(U,W_\alpha) \ge \underline{\lim}_{T\to\infty} \frac{1}{T} I_T(U,W_\alpha) > \overline{I}(\tilde{U},\tilde{W}_\alpha) - \frac{\alpha^2\epsilon}{4\pi}. \tag{6}$$

Since $\epsilon > 0$ is arbitrary, (6) means that

$$\overline{I}(U,W_\alpha) = \lim_{T\to\infty} \frac{1}{T} I_T(U,W_\alpha) = \frac{1}{4\pi}\int_{-\infty}^{\infty} \log\left[1 - \frac{\alpha^2 f_U}{f_V}\right]d\lambda,$$

implying $(\alpha U,V) \in \mathbf{A}$.

II. If f_W and f_Z are rational, then $\sigma(\lambda)$, $\sigma_+(\lambda)$ and $\sigma_-(\lambda)$ are continuous. Noting that $0 < \theta = \lim_{\lambda\to\infty} f_Z(\lambda)/f_W(\lambda) < \infty$, we can easily see that $\sigma(\lambda) = O(\lambda^{-2})$ as $\lambda \to \infty$. Consequently, $\sigma_-(\lambda) = O(\lambda^{-2})$. Thus the assumptions of (b) are satisfied and (21) holds.

Proof of Theorem 4: Let X be an arbitrary stationary process belonging to $\mathfrak{X}_W(P)$, and define τ and $\{\tau_n(T)\}$ as in the proof of Theorem 3. Since the first inequality of (58) is valid for all sufficiently small $\alpha > 0$, using Lemma 6 we obtain

$$\frac{1}{2\pi} \int \tau(\lambda)d\lambda \leq \varliminf_{T\to\infty} \frac{1}{T} \Sigma_n \tau_n(T) \leq P. \tag{60}$$

This means that $X \in \mathcal{S}_W(P)$.

APPENDIX

Proof of Lemma 3: Let us introduce some notations. Let S(P) be the class of all SDF's $f(\lambda)$ which satisfy

$$\int f(\lambda)d\lambda \leq 2\pi P/\theta$$

and $J(f)$ be a functional defined by

$$J(f) = \frac{1}{4\pi} \int \log\left[1 + \frac{f(\lambda)}{1+\sigma(\lambda)}\right]d\lambda.$$

Then, noting (10) and (15), we have

$$\Gamma(P) = \sup \{J(f);\ f \in S(P)\}. \tag{61}$$

We shall use the following inequalities,

$$\log(1+x) \leq x, \quad x > -1, \tag{62}$$

$$\log\left[1 + \frac{x}{z-y}\right] \geq \frac{x}{z}, \quad 0 \leq x \leq y < z. \tag{63}$$

Suppose first that $P \geq \frac{-\theta}{2\pi} \int \sigma\_(\lambda)d\lambda$. Denote by $\tilde{\Gamma}(P)$ the RHS of (16a). If $f \in S(P)$, then from (62)

$$4\pi J(f) = \int\left[\log\left[\frac{1}{1+\sigma_-(\lambda)}\right] + \log\left[\frac{1+\sigma(\lambda)+f(\lambda)}{1+\sigma_+(\lambda)}\right]\right]d\lambda$$

$$= \int\left[\log\left[\frac{1}{1+\sigma_-(\lambda)}\right] + \log\left[1 + \sigma_-(\lambda) + \frac{f(\lambda)}{1+\sigma_+(\lambda)}\right]\right]d\lambda$$

$$\leq \int\log\left[\frac{1}{1+\sigma_-(\lambda)}\right]d\lambda + \int\sigma_-(\lambda)d\lambda + \int f(\lambda)d\lambda$$

$$\leq \int\log\left[\frac{1}{1+\sigma_-(\lambda)}\right]d\lambda + \int\sigma_-(\lambda)d\lambda + 2\pi P/\theta = 4\pi\tilde{\Gamma}(P).$$

This inequality means that

$$\Gamma(P) \leq \tilde{\Gamma}(P). \tag{64}$$

From (9) and (10) we know that

$$\int[(1+\delta) - \min\{(1+\delta), (1+\sigma(\lambda))\}]d\lambda = \infty,$$

for every $\delta > 0$. Hence, for every fixed $\delta > 0$, there exists a function $u(\lambda)$ such that $0 \leq u(\lambda) \leq \delta$, $u(\lambda) = 0$ if $\sigma(\lambda) \geq \delta$, and

$$\int u(\lambda)d\lambda = \frac{2\pi P}{\theta} + \int\sigma_-(\lambda)d\lambda. \tag{65}$$

Define an SDF f by $f(\lambda) = u(\lambda) - \sigma_-(\lambda)$. Then $f \in S(P)$. Noting $u(\lambda)/(1+\sigma_+(\lambda)) \geq u(\lambda)/(1+\delta)$ and using (63) we obtain

$$4\pi J(f) = \int\left[\log\left[\frac{1}{1+\sigma_-(\lambda)}\right] + \log\left[1 + \frac{u(\lambda)}{1+\sigma_+(\lambda)}\right]\right]d\lambda$$

$$\geq \int\left[\log\left[\frac{1}{1+\sigma_-(\lambda)}\right] + \log\left[1 + \frac{u(\lambda)}{1 + \delta}\right]\right]d\lambda$$

$$\geq \int\left[\log\left[\frac{1}{1+\sigma_-(\lambda)}\right] + \frac{u(\lambda)}{1 + 2\delta}\right]d\lambda. \tag{66}$$

It follows from (61), (65) and (66) that

$$4\pi\Gamma(P) \geq \int \log \frac{1}{1+\sigma_-(\lambda)}d\lambda + \frac{1}{1+2\delta}\left[\frac{2\pi P}{\theta} + \int\sigma_-(\lambda)d\lambda\right]. \tag{67}$$

Since $\delta > 0$ is arbitrary, (67) gives

$$\Gamma(P) \geq \tilde{\Gamma}(P). \tag{68}$$

The equation (16a) follows from (64) and (68). Next, we assume that $P < -\frac{\theta}{2\pi}\int\sigma_-(\lambda)d\lambda$. Define sets $B_- \equiv B_-(P)$ and $B_+ \equiv B_+(P)$ in $\mathbb{R}$ by $B_- = \{\lambda;\ \sigma(\lambda) \leq A(P)\}$ and $B_+ = \{\lambda;\ \sigma(\lambda) > A(P)\}$. Note that $\sigma(\lambda) = \sigma_-(\lambda)$ for $\lambda \in B_-$. By (17) it is clear that a SDF f_0, defined by

$$f_0(\lambda) = \begin{cases} A(P) - \sigma_-(\lambda), & \lambda \in B_-, \\ 0, & \lambda \in B_+, \end{cases}$$

belongs to S(P). For any f in S(P),

$$\int f \leq \int f_0 = 2\pi P/\theta \tag{69}$$

and we wish to show that

$$\int \log\left[1 + \frac{f}{1+\sigma}\right] \leq \int \log\left[1 + \frac{f_0}{1+\sigma}\right]$$

or

$$\int \log\left[\frac{1+\sigma+f}{1+\sigma+f_0}\right] \leq 0.$$

But
$$\int \log\left[1 + \frac{f-f_0}{1+\sigma+f_0}\right] \leq \int\left[\frac{f-f_0}{1+\sigma+f_0}\right] = \int_{B_+}\frac{f-f_0}{1+\sigma} + \int_{B_-}\frac{f-f_0}{1+A(P)}$$

$$\leq \int_{B_+}\frac{f-f_0}{1+A(P)} + \int_{B_-}\frac{f-f_0}{1+A(P)} = \int \frac{f-f_0}{1+A(P)} \leq 0, \tag{70}$$

so that $4\pi J(f) \leq 4\pi J(f_0)$. It follows from (61) and (70) that $\Gamma(P) = J(f_0)$, yielding (16b).

<u>Proof of Lemma 4</u>: The relation (28) is clear from (27) and (14), and implies that S_T is negative definite. The assumption (21) means that $I_T(U, Q^{(\alpha)}) < \infty$ where $Q^{(\alpha)}(t) = \alpha U(t) + V(t)$. Thus S_T is a trace class operator by Lemma 1. Noting (11) - (13) we know that

$$\frac{f_U(\lambda)}{f_V(\lambda)} = -\sigma_-(\lambda) \le 1 - a_0, \quad \text{on } \{f_W \ne 0\}. \tag{71}$$

From (71), $R_{U,T} \le (1-a_0)R_{V,T}$, so $R_{V,T}^{-\frac{1}{2}} R_{U,T} R_{V,T}^{-\frac{1}{2}} \le (1-a_0)I$, or

$$S_T \ge (-1 + a_0)I, \tag{72}$$

where I is the identity operator. The relation (29) follows from (72).

<u>Proof of Lemma 6</u>: (a) At first we shall prove (35) by induction. We use the following inequality,

$$\left|\log(1+x) - \sum_{k=1}^{K-1} \frac{(-1)^{k-1}x^k}{k}\right| \le \frac{x^K}{K}, \quad 0 \le x \le 1, \quad K = 2,3,\dots . \tag{73}$$

From (33) we know that $(T^{-1})\Sigma_n \log(1+\alpha\tau_n(T))$ is bounded for sufficiently large T. Hence $(T^{-1})\Sigma_n\tau_n(T)$ and, consequently, $(T)^{-1}\Sigma_n\tau_n(T)^2$ are bounded. Therefore, there exists constants M_1 and T_1 such that

$$\frac{1}{T}\Sigma_n\tau_n(T)^2 \le M_1, \quad T \ge T_1. \tag{74}$$

For any fixed $\epsilon > 0$ there exists $\alpha_1 > 0$ such that

$$\begin{aligned} 0 &\le \int \tau(\lambda)d\lambda - \frac{1}{\alpha}\int \log(1+\alpha\tau(\lambda))d\lambda \\ &\le \frac{\alpha}{2}\int \tau(\lambda)^2 d\lambda < 2\pi\epsilon, \quad 0 < \alpha \le \alpha_1. \end{aligned} \tag{75}$$

Similarly, we know that there exists $\alpha_2 > 0$ such that

$$\begin{aligned} 0 &\le \frac{1}{T}\Sigma_n\tau_n(T) - \frac{1}{\alpha T}\Sigma_n \log(1+\alpha\tau_n(T)) \\ &\le \frac{\alpha}{2T}\Sigma_n\tau_n(T)^2 < \epsilon, \quad 0 < \alpha \le \alpha_2, \quad T \ge T_1. \end{aligned} \tag{76}$$

Fix a number $\alpha_3 \le \min(\alpha_0,\alpha_1,\alpha_2)$. It follows from (33) that

$$|(\alpha_3 T)^{-1}\Sigma_n \log(1+\alpha_3\tau_n(T)) - (2\pi\alpha_3)^{-1}\int \log(1+\alpha_3\tau(\lambda))d\lambda| < \epsilon, \qquad (77)$$

$$T \geq T_2,$$

where T_2 is a constant. From (75) - (77) we obtain

$$|\frac{1}{T}\Sigma_n\tau_n(T) - \frac{1}{2\pi}\int\tau(\lambda)d\lambda| < 3\epsilon, \quad T \geq \max(T_1,T_2),$$

yielding (35) for k = 1. We now assume that (35) is true for k = 1,...,K-1. Then,

$$\frac{1}{K}|\frac{1}{T}\Sigma_n\tau_n(T)^K - \frac{1}{2\pi}\int\tau(\lambda)^K d\lambda|$$

$$\leq |\frac{1}{\alpha^K}\sum_{k=1}^{K}\frac{(-1)^{k-1}\alpha^k}{k}(\frac{1}{T}\Sigma_n\tau_n(T)^k - \frac{1}{2\pi}\int\tau(\lambda)^k d\lambda)|$$

$$+ \frac{1}{\alpha^K}\left|\sum_{k=1}^{K-1}\frac{(-1)^{k-1}\alpha^k}{k}(\frac{1}{T}\Sigma_n\tau_n(T)^k - \frac{1}{2\pi}\int\tau(\lambda)^k d\lambda)\right|.$$

The second term on the RHS of the inequality converges to zero as $T \to \infty$, for all $\alpha > 0$, by the induction hypothesis. The first term is majorized by

$$\left|\frac{1}{T\alpha^K}\sum_{k=1}^{K}\frac{(-1)^{k-1}\alpha^k}{k}\Sigma_n\tau_n(T)^k - \frac{1}{T\alpha^K}\Sigma_n\log(1+\alpha\tau_n(T))\right|$$

$$+ \left|\frac{1}{2\pi\alpha^K}\sum_{k=1}^{K}\frac{(-1)^{k-1}\alpha^k}{k}\int\tau(\lambda)^k d\lambda - \frac{1}{2\pi\alpha^K}\int\log(1+\alpha\tau(\lambda))d\lambda\right|$$

$$+ \left|\frac{1}{T\alpha^K}\Sigma_n\log(1+\alpha\tau_n(T)) - \frac{1}{2\pi\alpha^K}\int\log(1+\alpha\tau(\lambda))d\lambda\right|.$$

The last of these three terms converges to zero as $T \to \infty$, for all $\alpha \leq \alpha_0$, by (33). Using (73), the first two are majorized by

$$\frac{\alpha}{T(K+1)}\Sigma_n\tau_n(T)^{K+1} + \frac{\alpha}{2\pi(K+1)}\int\tau(\lambda)^{K+1}d\lambda$$

$$\leq \frac{2M^{K-1}}{K+1}\frac{\alpha}{2T}\Sigma_n\tau_n(T)^2 + \frac{M^{K-1}}{K+1}\frac{\alpha}{2\pi}\int\tau(\lambda)^2 d\lambda \leq \frac{3M^{K-1}}{K+1}\epsilon$$

for $\alpha \le \min(\alpha_1, \alpha_2)$, by (75) and (76). This proves (35). We turn now to the proof of (34). For any fixed $\epsilon > 0$, there exist polynomials

$$P(x) = \sum_{k=1}^{K} a_k x^k, \quad Q(x) = \sum_{k=1}^{K} b_k x^k$$

such that $P(x) \le F(x) \le Q(x)$, $0 \le x \le M$, and $\int_0^M \{Q(\tau(x)) - P(\tau(x))\}dx < \epsilon$. It is clear that

$$\lim_{T\to\infty} T^{-1}\Sigma_n P(\tau_n(T)) \le \underline{\lim}_{T\to\infty} T^{-1}\Sigma_n F(\tau_n(T)) \le \overline{\lim}_{T\to\infty} T^{-1}\Sigma_n F(\tau_n(T)) \tag{78}$$

$$\le \lim_{T\to\infty} T^{-1}\Sigma_n Q(\tau_n(T)).$$

Using (35) we get

$$\lim_{T\to\infty} T^{-1}\Sigma_n P(\tau_n(T)) = (2\pi)^{-1}\int P(\tau(\lambda))d\lambda \le (2\pi)^{-1}\int F(\tau(\lambda))d\lambda - \epsilon, \tag{79}$$

$$\lim_{T\to\infty} T^{-1}\Sigma_n Q(\tau_n(T)) = (2\pi)^{-1}\int Q(\tau(\lambda))d\lambda \le (2\pi)^{-1}\int F(\tau(\lambda))d\lambda + \epsilon. \tag{80}$$

Since $\epsilon > 0$ is arbitrary, (34) follows from (78) - (80). Finally, (b) can be shown in the same manner as (a).

<u>Proof of Lemma 7</u>: Since $f_V(\lambda) = f_W(\lambda)(1+\sigma_+(\lambda)) \ge f_W(\lambda)$ for all λ in $\mathbb{R}$, we have $R_{V,T} \ge R_{W,T}$ and consequently $R_{V,T}^{-1} \le R_{W,T}^{-1}$. Therefore, $R_{X,T}^{\frac{1}{2}} R_{V,T}^{-1} R_{X,T}^{\frac{1}{2}} \le R_{X,T}^{\frac{1}{2}} R_{W,T}^{-1} R_{X,T}^{\frac{1}{2}}$, and finally we obtain

$$\text{Trace } R_{V,T}^{-\frac{1}{2}} R_{X,T} R_{V,T}^{-\frac{1}{2}} = \text{Trace } R_{X,T}^{\frac{1}{2}} R_{V,T}^{-1} R_{X,T}^{\frac{1}{2}} \le \text{Trace } R_{X,T}^{\frac{1}{2}} R_{W,T}^{-1} R_{X,T}^{\frac{1}{2}}$$

$$= \text{Trace } R_{W,T}^{-\frac{1}{2}} R_{X,T} R_{W,T}^{-\frac{1}{2}} .$$

The inequality (37) follows from (36) and (4).

Proof of Lemma 8: There exist $\Lambda > 0$ and $a > 0$ such that a rational function $\varphi(\lambda)$, defined by

$$\varphi(\lambda) = a(\Lambda^2 + \lambda^2)^{-1} \qquad (\Lambda = 3\pi\epsilon^{-1}),$$

satisfies

$$f(\lambda) \le \varphi(\lambda), \qquad |\lambda| \ge \Lambda.$$

Note that

$$\int_{-\infty}^{\infty} \varphi(\lambda)d\lambda = 3\pi\Lambda^{-1} = \frac{\epsilon}{3}.$$

Using Weierstrauss' theorem, we can show that there exists a rational SDF $\psi(\lambda)$ satisfying $f(\lambda) \le \psi(\lambda)$ for $|\lambda| \le \Lambda$, $0 < \psi(\lambda) \le f(\lambda) + \epsilon/(12\Lambda)$ for $|\lambda| \le 2\Lambda$ and $\int_{|\lambda| \ge 2\Lambda} \psi(\lambda) \le \epsilon/3$. We define the function $g(\lambda)$ by $g(\lambda) = \varphi(\lambda) + \psi(\lambda)$. Then we can show that $f(\lambda) \le \psi(\lambda) \le g(\lambda)$ for $|\lambda| \le \Lambda$, $f(\lambda) \le \varphi(\lambda) \le g(\lambda)$ for $|\lambda| \ge \Lambda$ and

$$\frac{1}{2}\int_{-\infty}^{\infty}[g(\lambda) - f(\lambda)]d\lambda$$
$$\le \int_0^{2\Lambda}[\psi(\lambda) - f(\lambda)]d\lambda + \int_{2\Lambda}^{\infty}\psi(\lambda)d\lambda + \int_0^{\infty}\varphi(\lambda)d\lambda \le \frac{\epsilon}{2}.$$

REFERENCES

[1] C.R. Baker, Capacity of the Gaussian channel without feedback, *Information and Control*, 37, pp. 70-89, 1978.

[2] C.R. Baker, Channel models and their capacity, in *Essays in Statistics: Contributions in Honor of Norman L. Johnson*, New York: North-Holland, 1983, pp. 1-16.

[3] C.R. Baker, Capacity of the mismatched Gaussian channel, *IEEE Trans. on Inform. Theory* (1987).

[4] R.G. Douglas, On majorization, factorization, and range inclusion of operators on Hilbert space, *Proc. Amer. Math. Soc.*, 17, pp. 413-415, 1966.

[4a] B. Eisenberg, Translating Gaussian processes, *Ann. Math. Stat.*, 41, 888-893 (1970).

[5] R.M. Fano, *Transmission of Information*, Cambridge: M.I.T. Press, and New York: Wiley, 1961.

[6] R.G. Gallager, *Information Theory and Reliable Communication*, New York: Wiley, 1968.

[7] M. Hitsuda and S. Ihara, Gaussian channels and the optimal coding, *J. Multivariate Anal.*, 5, pp. 106-118, 1975.

[8] I.A. Ibragimov and Yu. A. Rozanov, *Gaussian Random Processes*, New York: Springer-Verlag, 1978.

[9] S. Ihara, Mutual information in stationary channels with additive noise, *IEEE Trans. Inform. Theory*, IT-31, pp. 602-606, Sept. 1985.

[10] M.S. Pinsker, Computation of rate of generating information of stationary random processes and capacity of the stationary channel, *Dokl. Akad. Nauk SSSR*, 111, pp. 753-756, 1956 (in Russian).

[11] M.S. Pinsker, *Information and Information Stability of Random Variables and Processes*, San Francisco: Holden-Day, 1964.

[12] C.E. Shannon, Communication in the presence of noise, *Proc. IRE*, 37, pp. 10-21, 1949.